Representation recovers information

Article (Submitted Version)

Thornton, Chris (2009) Representation recovers information. Cognitive Science, 33 (8). pp. 1-30. ISSN 0364-0213

Representation Recovers Information

Chris Thornton
COGS/Informatics
University of Sussex
Brighton
BN1 9QH
UK
c.thornton@sussex.ac.uk

September 2, 2009

Abstract

Early agreement within cognitive science on the topic of representation has now given way to a combination of positions. Some question the significance of representation in cognition. Others continue to argue in favor, but the case has not been demonstrated in any formal way. The present paper sets out a framework in which the value of representation-use can be mathematically measured, albeit in a broadly sensory context rather than a specifically cognitive one. Key to the approach is the use of Bayesian networks for modeling the distal dimension of sensory processes. More relevant to cognitive science is the theoretical result obtained, which is that a certain type of representational architecture is *necessary* for achievement of sensory efficiency. While exhibiting few of the characteristics of traditional, symbolic encoding, this architecture corresponds quite closely to the forms of embedded representation now being explored in some embedded/embodied approaches. It becomes meaningful to view that type of representation-use as a form of information recovery. A formal basis then exists for viewing representation not so much as the substrate of reasoning and thought, but rather as a general medium for efficient, interpretive processing.

Keywords: representation, information, sensory embedding, close coupling, cognitive informatics

1 Introduction

Questions about representation-usage have been the subject of a lively debate in cognitive science for many years (e.g., Brooks, Reasoning, 1991; van Gelder, 1995; Beer, 2003; Svensson and Ziemke, 2005). The broad-based acceptance of the computational metaphor in the early years of the field went hand-in-hand with the idea that representation must be a key element of cognitive activity:

it is, after all, a key element of *computational* activity (Wheeler, 1996). But more recent times have seen mounting resistance to the idea (Winograd and Flores, 1986) with Brooks' publications on 'Intelligence without Representation' (Brooks, 1991b) and 'Intelligence without Reason' (Brooks, 1991a) creating a watershed for new thinking on the topic.

In developing this debate between pro- and anti-representationalist camps, cognitive science can be seen as having reworked the traditional dichotomy between rationalist and empiricist philosophies of knowledge. Anti-representationalism broadly aligns with empiricism, emphasizing the way cognition is grounded in agent/environment interaction. Representationalism broadly aligns with rationalism, emphasizing the way cognition is grounded in processes of reasoning. But in cognitive science the distinction has been made less anthropocentric and given a mechanistic foundation. Brooks' assertion that 'the world is its own best representation' thus contrasts with Locke's thesis 'there is nothing in the mind save that which is attained through the senses' primarily in its coverage of non-human intelligence and its provision of a more technical sense of constituency.

Cognitive science has also brought a stronger sense of there being middle ground in the debate. While dismissing the general value of representation in AI, Brooks, for example, was ready to include the operation of 'build maps' and even 'plan changes to the world' in his proposed subsumption architecture (Brooks, 1986). This suggested an acceptance of there being *some* role for representation. In the field of vision research, Ballard appeared to concede something similar. Questioning Marr's representation-based conception of visual processing (Marr, 1982), Ballard proposed instead the method of *animate vision* (Ballard, 1991). This would make full use of the possibilities of agent-action to facilitate the visual task.[1] But as with Brooks' situated approach, Ballard's animate-vision proposal exhibited some representationalist features, e.g., its recognition of the need for 'object centered reference frames' (p. 77).

While some implicitly conceded the possibility of middle ground, others argued for its inevitability. Accepting the viability of non-representational strategies in many cases, Kirsh (1996) noted using the world 'as its own best representation' must fail in those cases where salient features of the environment are not sensorily accessible. Clark and Toribio (1994) made a similar point in their definition of the 'representation-hungry' domain. This characterizes the situation where action must be coordinated with features of the environment which are not accessible through sensory experience.[2]

But while many authors have argued that the need for representation must 'kick in' at some point, the problem has been to give a formal account of when and *why*. This is the task tackled by the present paper. Using Bayesian networks to model distal-to-proximal structures, it shows how measurements can be made of distal sensory information, i.e., the content of proximal stimuli treated as information about distal phenomena. The constraint then emerges that sensory

[1] In the simplest case, the method exploits gaze control, described by Ballard as 'the central asset of animate vision' (Ballard, 1991, p. 61).

[2] Such domains are closely related to the ones which exhibit 'type 2' learning problems in the type-1/type-2 framework of (Clark and Thornton, 1997).

architectures must have the right representational properties in order to be informationally efficient. On that basis it is argued that representation use, in its most basic form, can be understood as a form of information recovery.

The paper is in eight sections. The next section (Section 2) introduces use of the Shannon information metric for measurement of distal sensory information. Section 3 introduces use of Bayesian networks for modeling distal-to-proximal mappings. Section 3 presents illustrative examples involving sensory schemes in autonomous and non-autonomous agents. Adressing the question of how efficiency can be achieved in distal sensing, Section 4 articulates the reversal principle for information optimization. This is related to cognitive concerns in Section 5, where the type of architecture required for sensory efficiency is shown to be a form of representational structure. Section 6 relates that form to other schemes from the literature, involving embedded representation. Section 7 considers how the provision of a theoretical grounding for such schemes affects their explanatory content. Section 8 offers concluding comments.

2 Distal sensory information

Informational analyses of sensing typically aim to measure the content of receptor signals, i.e., the informational properties of the patterns of energy impinging on receptors. In contrast, the present approach aims to measure the information content of states of the distal stimulus, as *mediated* by proximal stimuli. Concentration on the distal dimension places the analysis outside frameworks such as (Tononi *et al.* 1994; Tononi *et al.* 1996; Pfeifer and Scheier, 1999; Lungarella and Pfeifer, 2001; Lungarella *et al.* 2005; Lungarella and Sporns, 2005). These do not distinguish between distal and proximal stimuli. They also treat receptor signals as continuous data, to be rendered into a discrete form using first and second order density functions. The present approach, on the other hand, treats receptor data as discrete in origin. Other informational analyses of sensing are often technology-specific (e.g., Webster, 1999; Wilson, 2005; Usher and Keating, 1996) or modal in nature, concerning sensory processes of audition, vision, olfaction etc., (e.g., Sabins, 1978; Mather, 1999). Laming (1986) has developed an amodal analysis as have Brignell and White (1994) but these again focus on information relating to proximal stimuli.

The starting point for the present analysis is Shannon's definition of information (Shannon, 1948; Shannon and Weaver, 1949). This equates information with uncertainty-reduction. In Shannon's framework, the amount of information gained when probability attributions for a variable change, is the reduction obtained in the associated entropy (H).[3] Figure 1 illustrates the idea with regard to the weather. Treated as a variable whose possible values are known to

[3]The entropy H of probability distribution P is defined to be

$$-\sum_i P_i \log_2 P_i$$

where P_i is the probability of the ith alternative. Taking logs to base 2 allows information to be expressed in *bits*. Informally, entropy is a measure of the uniformity of the distribution.

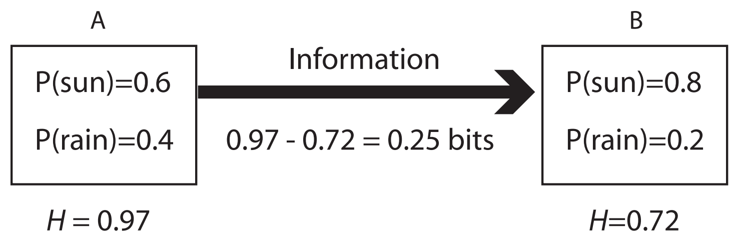

Figure 1: Information viewed as uncertainty reduction. The change from distribution A to distribution B implies a reduction in uncertainty, which can be measured in terms of the reduction in the entropy of the distributions.

be 'sun' and 'rain', the change from assumed distribution A to assumed distribution B implies a reduction in uncertainty, the size of which is found by subtracting the entropy of distribution B from the entropy of distribution A. Entropy being essentially a measure of the 'flatness' of the distribution, the difference in this case is the relatively small value of 0.25. Assuming logs taken to base 2, this yields an information value of 0.25 *bits*.

Typically, entropy-based measurement of uncertainty-change is used to establish the informational value of a message for a given receiver. The way in which the message alters the receiver's attributions of probability is established; measurement of the entropy-change then identifies the message's information content for that receiver. In the present case, however, the requirement is to measure the level of information that proximal stimulation provides about a *distal* stimulus. It will be necessary to adapt the measure to fit the purpose; but this should be done without undermining its generality. No arbitrary restrictions should be placed on the constitution of the distal stimulus — henceforth termed the 'target'. It should be possible for it to be a simple object (e.g., a bird), an object of multiple parts (e.g., a queue) or any salient feature of agent/environment engagement (e.g., a hiding place).[4] We also need to accommodate the possibility of proximal stimulation providing *misinformation* about the distal stimulus.

A strategy which seems promising initially involves using the *mutual information* measure (Cover and Thomas, 1991). We treat the target as one random discrete variable and the proximal stimulus as another. We then calculate the mutual information between the two. Letting T represent the target variable and S the proximal stimulus, we evaluate

$$I(S;T) = \sum_{t \in T} \sum_{s \in S} p(s,t) \log_2 \left(\frac{p(s,t)}{p_1(s) p_2(t)} \right). \tag{1}$$

[4]The paper is not concerned with the ontological issue of how or whether such stimuli might exist.

Given joint probability distribution $p(s,t)$ over variables S and T and marginal distributions $p_1(x)$ and $p_2(t)$, the mutual information $I(S;T)$ measures the amount of information variable S provides about variable T (and vice versa). This would seem to be what is wanted: a measure of the amount of information which the proximal stimulus provides about the target.

Unfortunately, the approach fails for two reasons. The amount of mutual information is not limited by the uncertainty associated with T itself. There is the potential for paradoxical 'measurements' in which sensory information exceeds the amount which would be obtained through direct access to the target. (Indeed, this will generally be the case if S has more values than T, since, other things being equal, a larger number of probabilities generates a larger entropy.) Second, there is the problem that levels of mutual information are always positive. There is no way to accommodate the way in which the proximal stimulus might be misinformative about the target. The problem is illustrated in Figure 2.

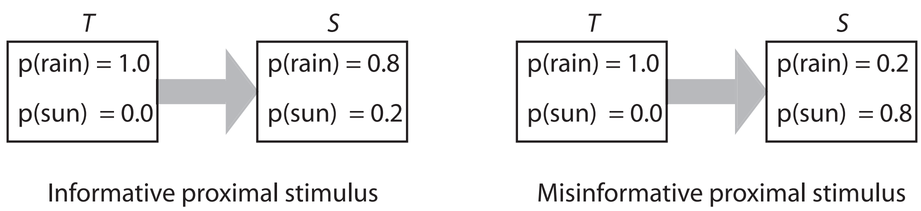

Figure 2: Informative v. misinformative sensing.

Mutual information measures the amount of information one variable provides about another but treats the two variables as separate entities with potentially different values and ranges. There is no sensitivity to the degree of match between *particular* values and probabilities: it is the relative entropic properties which count. For measurement of sensory information, however, we need precise alignment. Where one variable is treated as a proxy for the other, per-value correspondences are important and distinctions can be made between informative and misinformative scenarios (as in Figure 2).

We also require an approach which can register a difference between information and *mis*information, i.e., an information value with a boolean property. For this purpose, the *boolean bit* is introduced. This is a hybrid boolean/information value. The value of 1 boolean bit of information (termed 'bbit' below) is defined to be 1 bit of information where information increases certainty of a correct interpretation, and -1 bit where information increases certainty of an incorrect interpretation. A boolean bit is thus an ordinary bit of information but with a sign indicating whether or not the information contributes to a *correct* or *true* interpretation.

Measuring information in bbits, we have the means of balancing information against misinformation. For this purpose, the analysis will continue to view distal and proximal stimuli as instantiations of random discrete variables. The former will continue to be termed the 'target'. The latter will now generally be termed the 'sensor'. T_v will denote instantiation of the target to a particular value v, i.e., the distribution which assigns probability 1.0 to v and zero elsewhere. $S|T_v$ is then the conditional distribution on S given T_v.

The initial level of uncertainty regarding the target is the entropy of its maximally uncertain (i.e., flat) distribution. This can be calculated simply as the log of the number of values of T:

$$\log_2 |T| \tag{2}$$

In adopting S as a sensor for T, the agent treats conditional distributions on S as proxies for conditioning distributions of T. The change in the entropy of S when T adopts a certain value v is then

$$\Delta H(S, T_v) = \log_2 |T| - H(S|T_v) \tag{3}$$

This entropy-change, however, could be informative or misinformative. To decide which, we need to discover whether the conditional distribution $S|T_v$ forms a correct or incorrect approximation of T_v. To find the similarity of two distributions X and Y, we can use the Kullback-Leibler distance (Kullback, 1987).[5]

$$D_{KL}(X \parallel Y) = \sum_i X(i) \log \frac{X(i)}{Y(i)} \tag{4}$$

$S|T_v$ forms a correct approximation of T_v if it is more similar to T_v than to the distribution for some other state. Its Kullback-Leibler distance, in this case, must be less than half the distance for different states, i.e,. the distance for some T_v and T_w where $v \neq w$. We can therefore obtain a measurement in bbits using this half distance — labeled δ below — as a threshold. Specifically, we use

$$I(S, T_v) = \begin{cases} \Delta H(S, T_v) & \text{if } D_{KL}(S|T_v \parallel T_v) < \delta \\ -\Delta H(S, T_v) & \text{otherwise} \end{cases} \tag{5}$$

This defines the information content in bbits of a particular T_v given S used as a proxy. By averaging these quantities, we then arrive at an overall measure of the informational value of S with respect to T.

$$I(T{:}S) = \frac{\sum_{v \in T} I(S, T_v)}{|T|} \tag{6}$$

This is termed the *distal sensory information*. It can be viewed as the average information gain conferred by states of T given use of S. We can see it as measuring the bbits of information of those states treated as distal 'signals'. But it is also the bbits of information which the agent obtains simply by treating S as a sensor for T. On that basis, we might see $I(T{:}S)$ as the *distal sensor gain*:

[5] In this metric, the divisor distribution is assumed to have no zero values.

the gain obtained by treating S in this way. In what follows, both interpretations will be used although the term 'distal sensory information' will be preferred in general and the qualifier 'distal' will often be dropped since no non-distal values will be under consideration.

$I(T{:}S)$ performs in the way we need. Being expressed in bbits, the value may be positive or negative, indicative of information or misinformation. But we also have the the requisite cap on absolute value. Where S is generally informative about the state of T, the sensory information is positive. Where it is generally misinformative, the sensory information is negative. In neither case, however, can the absolute value exceed the uncertainty associated with T itself. There is no danger of paradoxically large measurements.

3 The distal-to-proximal mapping

Equation (6) provides the means of measuring the information provided by a proximal stimulus given some distal target. But to apply it, we must evaluate

$$H(S|T_v)$$

which entails accessing the conditional probability distribution $P(S = s|T_v)$. To discover what this is, we need to know how the target influences the sensor: we need to analyze the distal-to-proximal mapping.

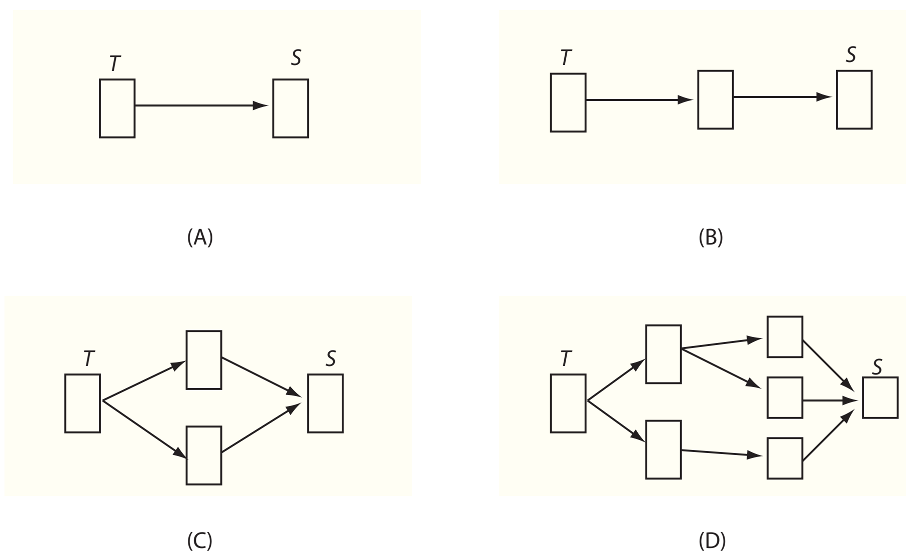

Figure 3: Sample sensory schemes.

There are many possible forms the mapping might take. We might have a trivial scenario in which the target influences the sensor directly. This might be the case, for example, where the target is something like ambient temperature and the sensor is a kind of temperature receptor. Here, there could be a 1-to-1 mapping between values of the target and values of the sensor. More commonly, the target will influence the sensor indirectly, via some configuration of intermediate variables. In such cases, determining the conditional probability distribution relating target to sensor involves taking the intermediate structure of influences into account.

Figure 3 illustrates some of the possibilities. Schematic (A) is the situation where target T influences sensor S directly. The other schematics illustrate cases with intermediate structure. Schematic (B) is relatively simple: the target influences *one* intermediate variable, which then influences the sensor. A real-world example would be the case where the target variable is the status of a gas ring on a cooker (alight or not) and the intermediate variable is ambient temperature. There may also be divergent/convergent influences between variables as illustrated in (C) and (D). In order to evaluate sensory information in a particular scheme, we need to discover how such influences *combine* to act on the sensor's distribution. This can be determined by treating the configuration of variables as a Bayesian network (Pearl, 1985; Pearl, 2000).

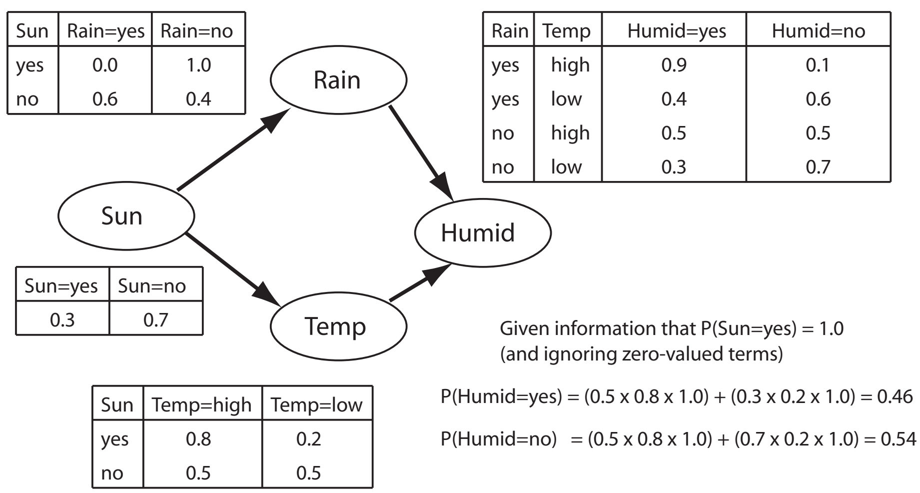

Figure 4: Bayesian inference using a network of four variables.

In the Bayesian network formalism, any variable Y which has a direct influence on variable X is said to be X's *parent*. Variable X is then said to be Y's *child* or *descendant*, with any child of a descendant also defined to be a descendant. The assumption is made that, in all cases, two variables are conditionally independent of all *non-descendants* given their parents. This is another way of saying that variables can be viewed as being influenced solely by their parents. A complete record of parental influences fully specifies the joint probability distribution for the variables. This 'complete record' is a Bayesian network.

Figure 4 illustrates use of a Bayesian network of four variables for purposes of calculating the relative probability of a sunny day being humid. Each variable in the network (the oval shapes) represents a different aspect of the weather (sun, rain, temperature, humidity). Each has an attached conditional probability table (CPT) which specifies the distributions imposed by possible combinations of parental values. The data are arranged so as to capture the idea that the state of the weather (sunny or not) influences the rainfall and the temperature, and that these two variables jointly influence whether it is humid. By propagating probabilities through the network, it is possible to calculate that the probability of a sunny day being humid is 0.46.

Within the Bayesian net formalism, there are general procedures for calculating distributions on variables, given constraints applying anywhere in the network (Korb and Nicholson, 2003). However, given the strict, feed-forward structure of this network, probabilities can be calculated by deriving joint probabilities working forwards (i.e., top-down) through the network as indicated in the lower, right part of the figure.[6]

Bayesian nets can also be used to assess the way relationships in a network of variables may produce distributions on a sensory variable, contingent on the adoption of a certain state of a target variable. Modeling the distal-to-proximal mapping as a Bayesian network then allows derivation of conditional distributions. This is exactly what we need for measurement of sensory information. To illustrate, the Bayesian network of Figure 4 is re-conceptualized as a sensory scheme in Figure 5. The 'Sun' variable is treated as target T and the 'Humid' variable becomes sensor S. 'Rain' and 'Temp' are then variables intermediate between target and sensor.

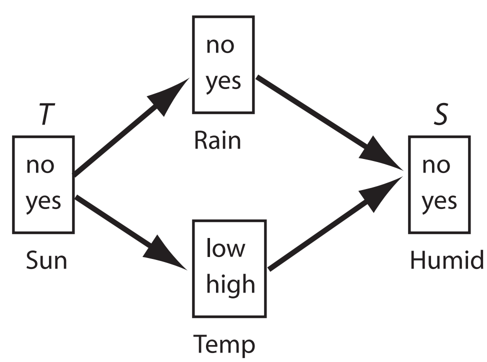

Figure 5: Bayesian net in a sensory configuration.

[6]The simplicity of the calculations in this case is due to the the way in which the zero probability for Rain=yes given Sun=yes, combined with the evidence that Sun=yes, ensures all terms associated with the top half of the Humidity probability table are zero-valued, and therefore not shown.

To measure distal sensory information we need to calculate probability distributions on the sensor given particular values of the target, as per equation (6). Given that the system of variables connecting target to sensor forms a feed-forward network, this can be done using forwards propagation of joint probabilities. By taking a Bayesian view of the network of influences connecting target to sensor, we are able to calculate distributions over sensory values consequent on distributions at the target. Invoking equations (4), (5) and (6) we can then calculate sensory information.

3.1 Example 1: the green-ball robot

Using this Bayesian approach, distal-to-proximal mappings of any complexity can be modeled in a form which enables distal sensory information to be measured. For a simple illustration, consider Figure 6. This depicts the functionality of a humanoid toy robot whose behavioral repertoire includes 'bowling' a small, green ball at some red skittles.[7] The robot uses a green-light sensor to discriminate the presence of the ball in its 'visual field'. On detecting the ball (i.e., on registering green light), a ball-grasping action is triggered and this leads on to

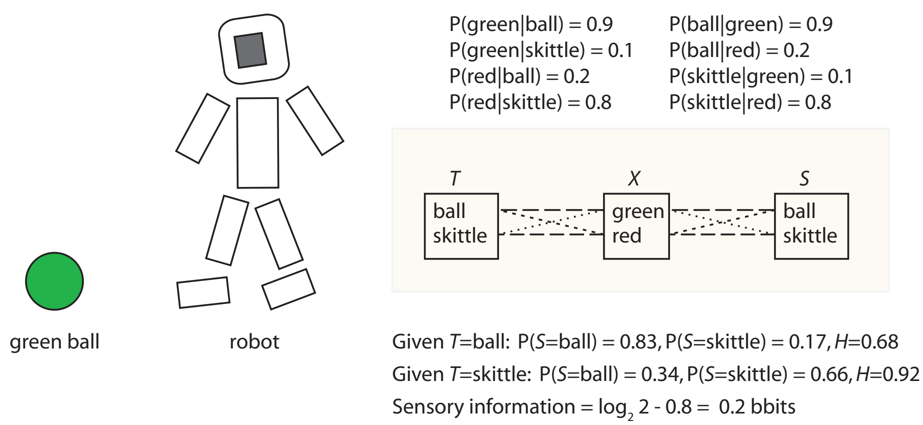

Figure 6: Measurement of distal sensor information.

the launching of the ball in the forwards direction.

The figure presents a model of the sensing utilized and evaluates sensory performance. The distal stimulus is taken to be the discriminated object. Possible values for T and S are thus 'ball' and 'skittle'. All conditional relationships are assumed to operate on a left-to-right basis, i.e., every dashed line has an implicit arrow pointing to the right.[8] Target and sensor are not directly linked, however. The target influences the sensor via intermediate variable X, which is the color of the object.

[7] The example is based on the Robosapiens 2 toy robot manufactured by Wowee Robotics.
[8] This applies to all the ensuing schematics.

Inter-variable influences (conditional probabilities) are as indicated in the upper part of the figure. These are also represented schematically using dashed lines. For each conditional probability there is a line connecting the conditioned value (on the right) to its conditioning value (on the left). Longer dashes indicate higher probabilities. Thus the conditional probability of X=green given T=ball is 0.9 (the connecting line uses a long dash) while the conditional probability of X=red given T=ball is 0.2 (the connecting line uses a short dash). The general situation is that balls are typically green and skittles typically red.

Applying forwards propagation, we can derive conditional probability distributions for S and calculate the information obtained for any particular state of the target. On that basis, we can measure sensory information. The steps in the procedure are summarized in the lower part of the figure. The average entropy for generated distributions on S is 0.8. Given that the maximum information value for this variable is $\log_2 2 = 1$, we then obtain a sensory information of 0.2 bbits. This sensing scheme thus shows a level of efficiency precisely one fifth of the optimum.

3.2 Example 2: *Chrysopa septempunctata*

The method can also be applied to models of natural sensory systems. For an illustration, consider Figure 7. This depicts distal sensing in the Green Lacewing

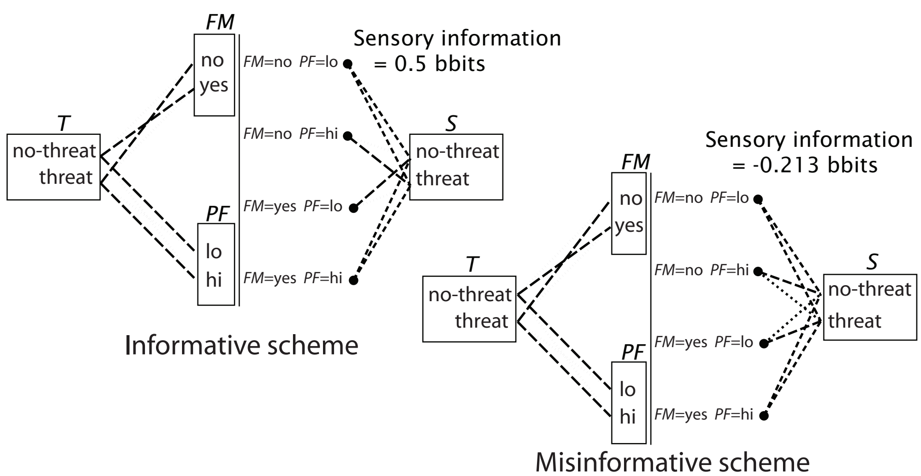

Figure 7: Capture-threat sensing in the green lacewing moth.

moth *Chrysopa septempunctata*. These insects have evolved a remarkable tactic for evading capture by predatory bats in the final stages of attack. Making use of sensory information regarding the frequency of echolocation pulses and the presence or absence of a frequency-modulation (FM) component in the call, these creatures can detect the threat of imminent capture and 'fold their wings and nosedive out of the sky' (Smith, 2000, p. 81).

Figure 7 contrasts two sensing schemes for this response: an informative scheme (upper, left) and a misinformative scheme (lower, right). Dashed lines are used to represent probability as before. However, here, we have the complication that values of the sensory variable are conditional on *combinations* of intermediate values. To handle this schematically, nodes are introduced (the small, filled circles) to represent combinations of intermediate values (shown immediately to the left).

The target is threat of attack. This influences the sensor via two intermediate variables, one of which (*FM*) is the presence or absence of an FM component in the echolocation call, and the other of which (*PF*) is the frequency of emitted pulses. Here and elsewhere 'lo' = low and 'hi' = high.[9]

Sensory information in the informative scheme is 0.5 bbits. (A listing of the conditional probabilities used is given in Appendix 1.) In contrast, the misinformative scheme depicted in the lower, right part of the figure produces a negative gain of -0.213 bbits, i.e., just under a third of a negative bbit. Intuitively, this is because the *FM*=no, *PF*=hi combination is correctly associated with *S*=threat in the informative structure but not in the misinformative structure.

3.3 Example 3: Electronic stability sensing

While the measurement framework can be applied to models of natural and autonomous agents, it is not limited to that usage. It can be applied to any

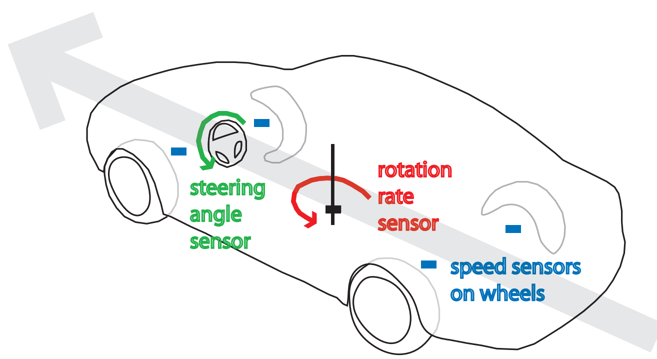

Figure 8: Sensing for electronic stability control.

distal sensing scheme where the distal-to-proximal mapping can be modeled as a Bayesian network. The framework can thus be applied to sensing schemes used in non-autonomous systems.

Consider the electronic stability system of the modern motorcar. An evolution of the automatic braking systems (ABS) from the 1980s, these more recent systems utilize multiple sensors and computer-controlled moderation of braking

[9] In reality, pulse frequency can range from 7 pulses per second to over 30 pulses.

for purposes of skid-avoidance. In a simple case, a sensing system might make use of inputs from a yaw sensor and wheel-speed sensors (cf. Figure 8) to control understeer — the process whereby a car fails to turn right or left due to skidding forwards.

Figure 9 presents a model in which the target is taken to be the trajectory of the vehicle. Possible values are 'oversteer', 'understeer' and 'balanced'. In

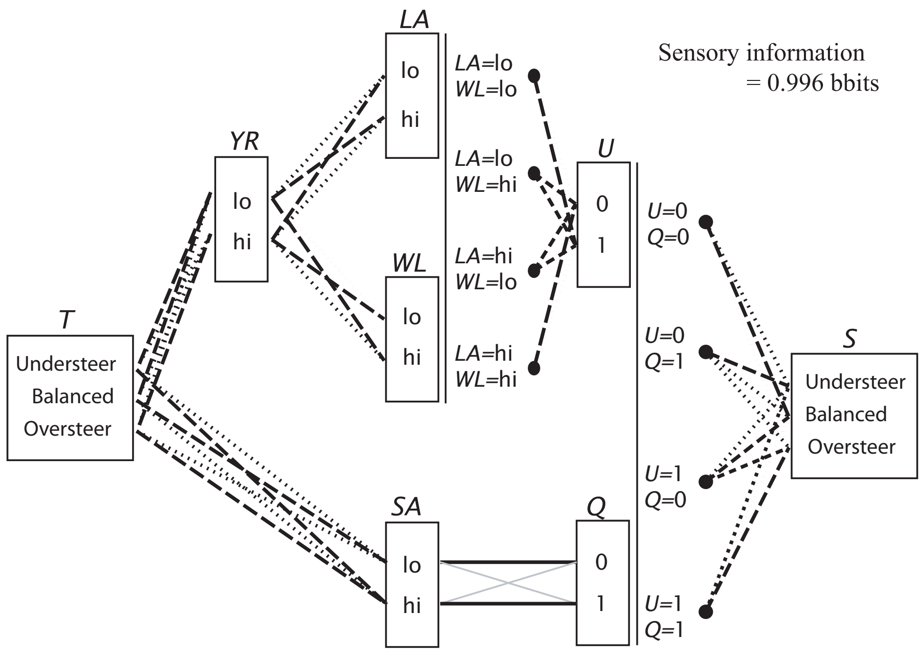

Figure 9: Informational model of an ES system.

this model, the target directly influences two intermediate variables: the current steering angle and the yaw rate of the vehicle — the rate at which it is rotating around its central point. The latter in turn influences lateral acceleration (sideways motion) and wheel-lock. There are then two further variables (U and Q) which jointly influence the sensory variable.

Given the indicated conditional probabilities, the sensory information is 0.996 bbits, compared to a maximum of 1.538 bbits. The system exhibits a relatively high level of efficiency. The example also illustrates that variables in the distal-to-proximal mapping need not correspond to physical objects or objective properties of the agent's environment. As in this case, they may correspond to abstract properties of the agent, or to properties of the agent/environment interaction.

4 Achievement of sensory efficiency

Attention can now turn to the main question for the approach, namely What properties of a distal sensing scheme most affect its informational efficiency? For purposes of answering the question, a new type of variable is introduced. Any variable which can be accessed but not influenced by the agent is now defined to be a *receptor variable*. Receptor variables can thus be viewed as the channels of proximal stimulation. By specifying a mapping between a target and

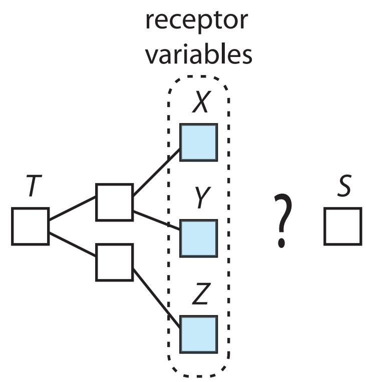

Figure 10: A distal sensing problem.

a set of receptor variables, it is then possible to define a distal sensing *problem*, as in Figure 10. Finding the solution to such a problem involves determining what structure of variables interposed between receptors (shaded boxes in this and ensuing schematics) and sensor will yield maximum sensory information. In terms of Figure 10, the arrangement to be determined is the one which maximizes information on the basis of T's (indirect) influence on X, Y and Z.

What, then, are the factors influencing efficiency? To start, we focus on the simple, direct-mapping scenario of Figure 11 (A). Here, target variable T is itself defined to be a receptor. As we might expect, the configuration which maximizes informational efficiency is then very simple. To maximize gain we must ensure conditional probabilities are configured so that each distinct value of the target is conditionally associated (with maximum probability) with its counterpart in the sensor. This is the effect implemented by the 1-to-1 mapping.

Taking a small step in complexity brings us to scenario (B). Here we have separate target and receptor variables and a conditional distribution in which the upper value of the receptor variable (X) is associated equally strongly with the two upper values of the target. For maximum sensory information, in this case, conditional probabilities connecting the receptor to the sensor must be configured in a specific way. If T's influence on X ensures that X has the value X_1 *or* X_2 with equal probability whenever T has the value T_1, we need to configure conditional probabilities affecting S so that X_1 and X_2 produce $S = T_1$

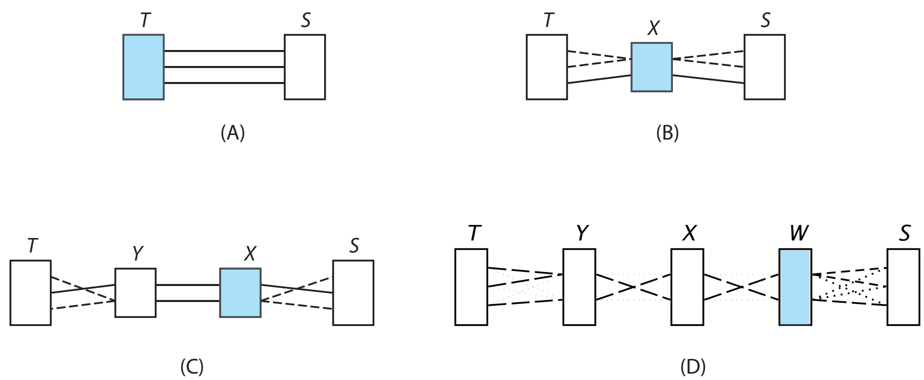

Figure 11: Simple sensing scenarios.

with equal probability. In general, we need conditional probabilities between X and S to *mirror* those between T and X. Sensory information is maximized by ensuring that S produces the best possible replication of the behavior of T; this is achieved by mirroring the relevant influences on any intermediate variable(s).

The situation stays essentially the same if we have a chain of intermediate variables connecting target to sensor, as in schematics (C) and (D). Where the receptor variable impacts the sensor directly (as it does in both these scenarios), the final link in the chain (i.e., those influences which act directly on the sensor) must be configured to mirror the influences acting on that variable. But in this case it is the cumulative influences which are involved and the mirroring effect may not be visually apparent, as exemplified in (D).

4.1 Optimization using inter-variable reversal

More challenging situations arise if we allow for the possibility of distribution across variables, i.e., situations in which one variable influences more than one child. The simplest case is depicted in the left-hand schematic of Figure 12. In addition to the target and sensor, this scenario features two receptor variables, both of which are directly influenced by the target. Information about the target is *distributed* across the two intermediate variables. When taken in combination, however, their values are mutually constraining with regard to their origin in T.

The only possible situation which might lead to $X = 1$ and $Y = 0$ is $T = 1$. Taking this into account allows the underlying value of T to be determined with full confidence. Maximization of sensory information is achieved by making $S = 1$ conditional on $X = 1$ and $Y = 0$ *jointly*. This is the effect achieved in the illustrated configuration. The unbroken line connecting to $S = 1$ shows this binding results with probability 1.0 if $X = 1$ and $Y = 0$. The other two unbroken lines mirror other 1-to-1 correspondences.

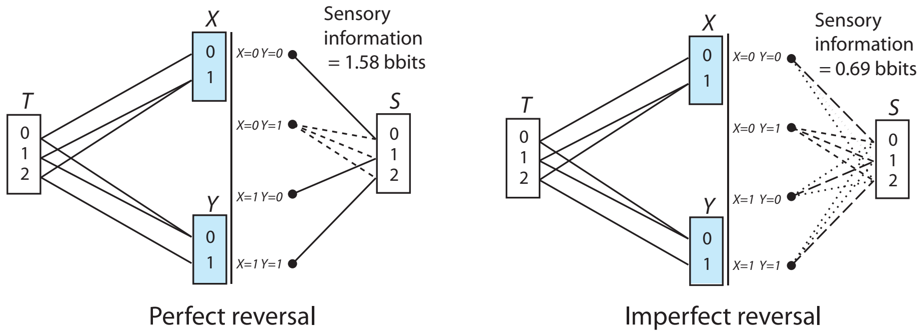

Figure 12: Perfect and imperfect reversal.

In this case we see how maximization of sensory information requires *reversing* the distributive effect applying to the value of T. This is similar to the requirement in the case of intra-variable distribution. But here the distribution operates *across* variables. Reversing it requires use of a correctly configured multi-parented variable — an *encapsulative* variable as it will be called. The conditional probabilities for this must be such as to capture and exploit the mutual constraints created in the distribution.

In sum, reversing distributive effects where multiple variables are involved involves use of matching encapsulative variables whose conditional relationships mirror the associations which produced the distribution. We need the distributive 'fan out' to be mirrored by the encapsulative 'fan in'. If there is no encapsulative variable, or if the conditional relationships do not produce the required reversal, sensory information is degraded. As an illustration consider the right-hand schematic of Figure 12. This is a version of the scheme depicted on the left but using non-reversing conditional probabilities. The result is an information value well below the optimum of 1.58 bbits.

4.2 The reversal principle

Distal sensing tasks can vary in several ways. But formalized in terms of target/sensor/receptor variables they can be analyzed in a way that reveals the requirements for informational efficiency. The general principles of optimization then come to light. Where the scenario exhibits perfect 1-to-1 correspondences between target and receptor values, as in Figure 11 (A), the sensory task is trivial. But where there is distribution of target values across receptor values, e.g., Figure 11 (B, C and D), the latter cannot unambiguously identify the former. Rather they identify them probabilistically and in this case, optimization is achieved by ensuring all intra-variable, distributive effects are fully reversed.

Where these is divergence in pre-receptor influences, we have distribution of information across multiple variables. Implicit, mutual constraints are then possible, in which case optimization involves another kind of reversal. Here it is the inter-variable distributive effects which must be reversed through the medium of encapsulation. The principle for optimizing sensor gain is thus to ensure that *all* distributive effects are properly reversed. Stated more precisely,

> *distal sensory information is maximized when there is matching reversal for all distributive effects*

This is the *reversal principle* for distal sensing. Contained within it is a structural requirement. Where distributive effects operate across variables, it is essential that encapsulative variables have access to (i.e., are influenced by) relevant source variables. This means the encapsulative sub-structure must be reversely isomorphic with the distributive sub-structure. A structural correspondence is entailed. This does not mean there has to be a 'visual match' between the two parts of the configuration, however. It only means those aspects of the structure which have distributive properties must be matched by variables with corresponding encapsulative properties.

5 Encapsulation viewed as representation

The technical content of the paper has now been set out. Sensory information relating to a distal stimulus has been shown to be measurable using Bayesian-network models of distal-to-proximal structure. The achievement of sensory efficiency has been shown to depend on the degree to which distributive effects are properly reversed. The general conclusion is that sensory schemes must exhibit reversal in order to be informationally efficient.

While informational analysis of sensing has a long tradition, (e.g., Barlow, 1961; van Hateren, 1992; Tononi *et al.* 1994; Tononi *et al.* 1996; Pfeifer and Scheier, 1999; Lungarella and Pfeifer, 2001; Hillis *et al.* 2002), previous approaches have not probed the distal dimension of the process in the way the present framework does. For purposes of sensory engineering, it has the advantage of formally accommodating the informational ramifications of environmental structure. But the more relevant benefit, for present purposes, is located in its *cognitive* implications. These emerge from a consideration of the representational properties of encapsulative structure.

At an abstract level, encapsulation may be viewed as the extraction of a particular *invariant*. Indeed, it is the defining property of an encapsulative variable to *reproduce* that invariant value of the parent from which the distributed pattern emerges. This has implications for the way we interpret the function of such variables. Where one encapsulative variable operates under the influence of the values of some other encapsulative variable (cf. Figure 9), reversal is fulfilled specifically on the basis that the latter properly reproduces values of the distributive counterpart. Values of the utilized encapsulative variable then play the role of informational surrogates *within* the scheme. A reversing encapsulative structure fulfilling a sensory function is constituted of encapsulative variables which the scheme itself treats as informational surrogates for distributive variables.

Consider Figure 13. In this scenario, the target is taken to be a certain property of a surface, namely its ability to provide the sensory agent with 'support'. (The possible values are taken to be 'supports', 'flexes' and 'gives-way'.) Sensing is mediated by three receptor variables. The first is the behavior of the agent, assumed to be either 'mob' (mobile) or 'sta' (static). The second is relative size ('rel-size'), i.e., the size of the agent relative to the surface, with possible values being 'sma' (small) and 'lrg' (large). The third is relative extension ('rel-extn'), i.e,. extension of the surface relative to the agent, with possible values 'hor' (horizontal) and ' ver ' (vertical).[10]

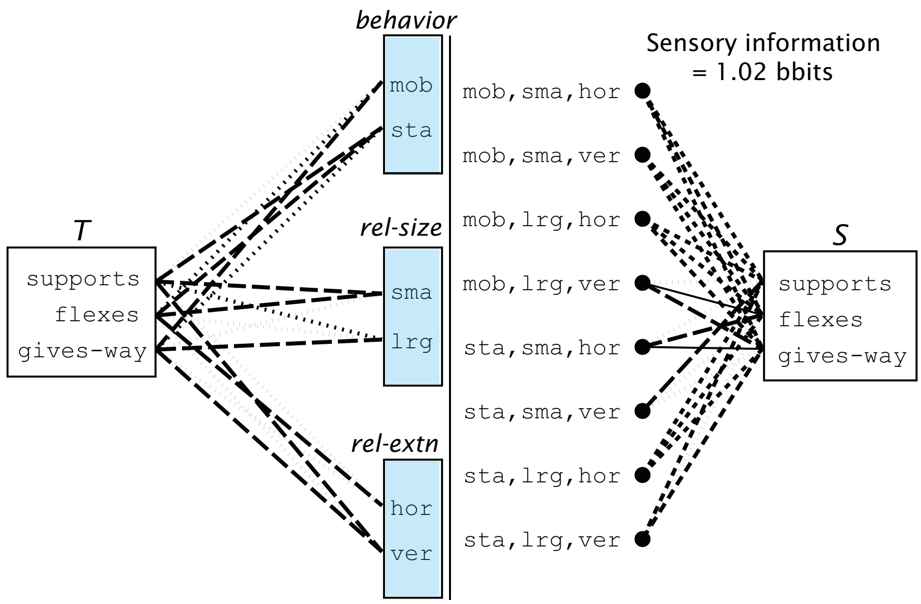

Figure 13: Encapsulative surrogacies.

Sensor S provides perfect reversing encapsulation with regard to T. Sensory information is 1.02 bbits and S can legitimately be viewed as a vehicle for reproducing values of T. Were it to feature in any further encapsulative structure, it would then play the role of an information surrogate for that property. In this case, the encapsulative variable serves as a proxy for a property of agent/environment interaction. But there is also the potential for surrogacies to bridge the agent/environment divide altogether. In a case where the target can be viewed as being exclusively part of the environment, the encapsulative variable then forms form an *internal* surrogate for an *external* entity.

Serving to link properties, features and entities of both agent and environment, a hierarchy of encapsulative variables forms a structure of surrogacies

[10]The example is loosely based on Gibson's supporting surface example (Gibson, 1979, p. 127).

whose efficient operation depends on its mediation of representational activity. An encapsulative structure is thus a *kind* of representational structure. But what kind exactly? On a number of counts, it departs from the familiar case of classical, symbolic representation. In particular,

- **There is no computer**: the representational structure does not presuppose the functioning of any kind of supervisory computational process.
- **There is no symbol processing**: the representational structure is not mediated by sequential processing of any form.
- **There is no world model**: the structure is not defined in terms of any 'external' universe.
- **There is no reasoner**: the system does not perform any kind of inference or deduction.
- **There is no encoding**: the representational structure is not stated or encoded in any kind of language.
- **There is no subject-object distinction**: encapsulative structure implies no particular division between agent and environment.

The type of representational activity emerging from the analysis thus differs in a range of ways from those familiar forms based on symbolic encoding, class hierarchies and relational databases (Brachman and Levesque, 1985). On the

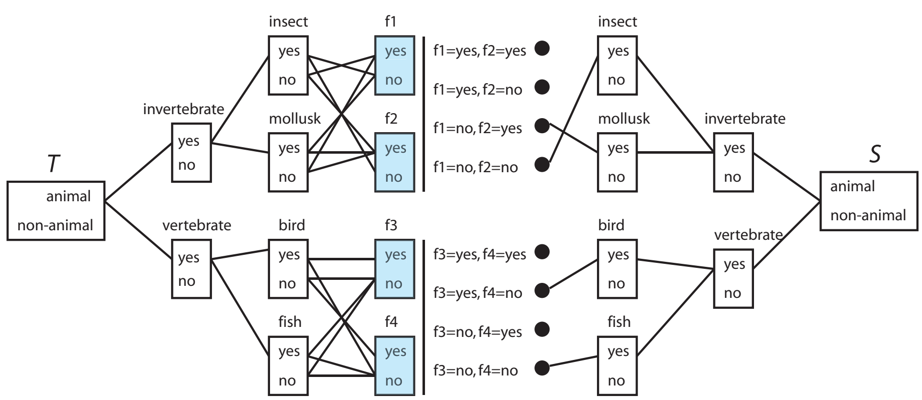

Figure 14: Encapsulative mediation of class subsumption.

other hand, being grounded in the formalism of Bayesian networks,[11] it has the power to *model* applications which feature in those forms. Consider the case of the generalization hierarchy. This representational construct can be modeled by

[11] As probabilistic graph models, Bayesian nets can model any dynamical or computational system.

an encapsulative structure through the expedient of letting multiple surrogacies implement class subsumption. Figure 15 illustrates how this might work in a class hierarchy for animals.

In this model, target T is an object classified either as 'animal' or 'non-animal'. Variable S is the sensor while f1, f2, f3 and f4 are receptors. Perfect reversal is achieved through the medium of matching encapsulative to distributive structure as usual. This also guarantees that the class-containment relations shaping the distributive structure are reproduced in encapsulative relations. An implicit model of class-subsumption is thus obtained.

Encapsulative representation can also model symbol-processing applications where conclusions are derived though reasoning. In such cases, the encapsulative system must use surrogacies to model relations, and propagation to model inference. Figure 15 illustrates the effect in an example involving kinship relations.

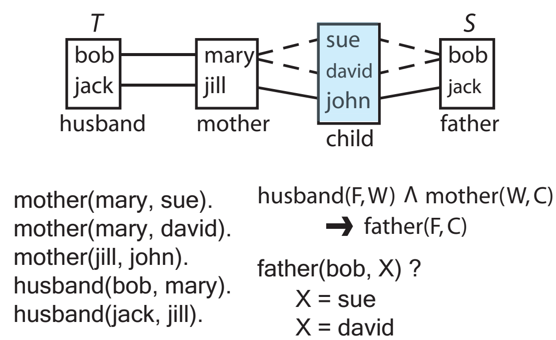

Figure 15: Encapsulative mediation of symbolic logic.

In the lower part of the figure, we see a traditional, deductive framework for the kinship domain made up of some facts, e.g., mother(mary, sue), and a rule which allows a 'father' relationship to be deduced on the basis of a 'mother' and a 'wife' relation. Use of this rule enables us to determine that 'bob' is the father of the two individuals 'sue' and 'david'.

In the upper part of the figure we see how this inferential activity might be modeled by a distributive/encapsulative structure. The identify of the father is here treated as variable S, i.e., as proxy for the identity of the husband. Given the evaluation husband=bob, propagation of values will correctly instantiate S, while giving equal probability to child=sue and child=david. The encapsulative structure mimics the behavior of the rule system in using the husband/mother dependencies for purposes of instantiating the 'father' variable. The inferential effect is then captured in the giving of equal probability to 'sue' and 'david' as the mediating value of 'child'.

6 Related representation concepts

While the notion of representation emerging from the analysis differs in various respects from GOFAI encoding, it has features in common with some of the novel, non-symbolic conceptions of representation being explored in embedded/embodied approaches. Such approaches tend to move away from traditional notions of symbol processing (e.g., Fodor, 1975; Newell and Simon, 1976; Newell and Simon, 1972; Pylyshyn, 1987), towards conceptions of representation which emphasize surrogacy, sensory embedding and system-functional content.[12]

Consider, for example, the view of Harnad (1990). His non-symbolic concept of representation explicitly stressed the idea of constructs being sensorily grounded. He suggested 'the symbolic functions would emerge ... as a consequence of the bottom-up grounding of categories' names in their sensory representations.' (Harnad, 1990, p. 335)

From the same era, Haugeland's definition of representation placed strong emphasis on surrogacy relations.

> 'We will reserve the term 'representation' for those stand-ins that function in virtue of a general *representational scheme* such that: (a), a variety of possible contents can be represented by a corresponding variety of possible representations; (b) what any given representation (item, pattern, state, event, etc.) represents is determined in some consistent or systematic way by the scheme; and (c) there are proper (and improper) ways of producing, maintaining, modifying, and/or using the various representations under various environmental and other conditions.' (Haugeland, 1991, p. 62)

Also emphasizing the fundamental nature of surrogacy is the definition from (Haselager *et al.* 2003). This defines an internal representation as 'an identifiable inner state within a system that stands in for another (internal or external) state and that as such plays a causal role for (or is used by) the system generating its behavior' (Haselager *et al.* 2003), cf. (Smith, 1996). Clark has also stressed the representational significance of 'inner surrogates' in a range of publications (Clark and Toribio, 1994; Clark, 1996; Clark and Thornton, 1997; Wheeler and Clark, 1999; Clark, 1997; Clark and Grush, 1999; Clark, 2003; Clark, 2008). Clark and Toribio's early framing of the 'representation-hungry domain' (Clark and Toribio, 1994, p. 31) treated sensorily-embedded representation as one of *two* significant cases, the other being more akin to symbolic representation. But recently Clark and others have gone further in stressing the specifically 'sensory' case. Indeed, Clark has more recently abandoned 'the idea of an executive center where the brain carries out high-level reasoning (Clark, 1997, p. xiii), which would seem to prioritize the 'sensory' part of the original definition.

[12]In Clark and Grush's view (1999, p. 9), the move is towards notions which 'tie the idea of internal representations to ideas involving tracking the distal, the absent, and the non-existent.'

Clark's work also moves deliberately away from conceptions of complete and consistent world models. In their stead, Clark envisages a 'large but fragmentary suite of internal representations' (Clark, 2008, p. 146), embedded structures which serve to 'refocus' domains of engagement. Such structures are often envisaged to be characteristically sensory in nature. Frequently cited, for example, is the case of emulator circuits (Clark and Grush, 1999; Kawato *et al.* 1987; Dean *et al.* 1994). These are mechanisms which 'replicate certain aspects of the temporal dynamics of the larger systems' (Clark, 1997, p. 23) but whose representational value is seen as originating in the way they mediate surrogacies constitutive of virtual *sensory* feedback (Clark and Grush, 1999; Grush, 2004).[13]

Concepts of sensorily-embedded representation steer clear of the subject-object distinction which plagues classical symbolic representation (Varella *et al.* 1991; Bickhard and Terveen, 1995). They are better able to accommodate an enactive or ecological view, in which representational content is associated with the possibilities of agent/environment interaction (Noë, 2004). In Bickhard and Terveen's 'interactivist' proposal (Bickhard and Terveen, 1995), for example, 'an encoding's having representational content is a property of the functional usage of the encoding by the system — it is a property of the system knowing what the encoding is supposed to represent — *and not a property of the encoding element itself*' (p. 57, original emphasis). This is also strongly emphasized in encapsulative representation, where it is the functional usage of a value which fixes its content. Bickhard and Terveen also stress the role of *mis*information, treating it as a requirement that a representation must 'involve some sense of *use* that can be wrong, and representation must be capable of being wrong for the system itself' (Bickhard and Terveen, 1995, p. 57). Again this is reflected in encapsulative representation, where variable evaluations can be 'wrong' (i.e., misinformative) for the embedding sensory scheme.

Many other researchers are developing notions of representation which emphasize surrogacy, sensory-embedding and system-functional content. Barsalou, for example, emphasizes all three factors in a framework which draws together cognition with perception.[14] His theory of 'perceptual symbol systems' (Barsalou, 1999) aims to 'demonstrate that it is possible to ground a fully functional conceptual system in sensory-motor mechanisms' (Barsalou, 1999, p. 608). In a similar vein, Prinz has developed a notion of conceptual representation built around 'proxytypes' (Prinz, 2002). The central idea here is that 'concepts are internally structured detection mechanisms' (Prinz, 2002, p. 125). In other words, Prinz proposes that internal representations of concepts can be understood in quasi-sensory terms. Also significant for the way they explore the connection between sensing and representation are the approaches of Gärdenfors (2000), Hesslow (2002) and Grush (2004).

[13] Hesslow, on the other hand, adopts an anti-representational position in advancing a similar proposal (Hesslow, 2002).

[14] Like Gibson (1979), Barsalou seems to have a radical application of Occam's Razor in mind with regard to notions of cognitive 'level'. In his view 'where we actually draw the line between perception and cognition may not be all that important, useful or meaningful,' (Barsalou, 1999, p. 589).

7 Explanatory implications

The concept of 'representational activity' emerging from the analysis dispenses with many of the commitments of classical, symbolic representation, including the computer, the world model, the reasoner, the encoding and the subject/object distinction. Like several of the embedded/embodied approaches cited, it emphasizes surrogacy, sensory embedding and system-functional content. Unlike them, it relies on theoretical task analysis rather than on analysis of empirical evidence. Does this make any difference to the explanatory account engendered? There are in fact a few points of detail where task-based analysis can serve to focus or extend empirically-based explanation.

What is representation?

The difficulties involved in answering this key, ontological question are well documented (Dreyfus, 1979; Searle, 1980). For some, any account constructed in terms of the concept of symbolic encoding is tautological (Bateson, 2000). Others identify an intrinsic circularity in the practice of symbolic representation which means the question cannot be coherently answered in that context (Harnad, 1990; Stewart, 1995). Bickhard and Terveen note there is no way in this paradigm (they call it 'encodingism') that 'representation can emerge out of phenomena that are not themselves already representational' (Bickhard and Terveen, 1995, p. 76). On that basis, there is no way of accounting for what representation really *is* in terms of description languages and symbolic encoding.

To the extent that approaches adopting an embedded/embodied perspective do away with such commitments, they promise to surmount problems of circularity. But the task analysis confirms this does not necessitate throwing out the concept of 'symbol' altogether. Rather it meshes with the idea that symbols may emerge as functional constructs. Clancey expresses it nicely when he notes (Clancey, 1997) that as a result of the 'complex interplay between ... inside and outside ... classification couplings may come to *function* as symbols in inferential reasoning' (p. 11). His emphasis that the interplay operates at 'many levels of organization simultaneously' (p. 75) further echoes the functionality of encapsulative representation. In terms of this *functional* conception of what a symbol really is, representation may continue to be seen as symbol-mediated.[15]

The task analysis meshes equally well with Clark and Grush's conception of representation. In (Clark and Grush, 1999), 'fullblooded internal representation' is defined to be 'specific states and/or processes whose functional role is to act as de-coupleable surrogates for specifiable (usually extra-neural) states of affairs' (Clark and Grush, 1999, p. 8). Rather than being imposed from the 'outside', symbols are envisaged as emerging as functional entities within the system. The sense of symbol-usage arising from the task analysis is similar. It also accommodates situations where distributive effects are grounded in agent/environment interaction. This is particularly illustrated in the model of

[15]As Clancey comments, 'researchers who held tenaciously to the idea of *symbols* in the brain were justified in doing so, but the physical nature, development, and reconstructive aspect of those symbol *systems* is quite unlike the labels and their manipulation in descriptive models.' (Clancey, 1997, p. 11).

electronic stability sensing (Figure 9), where several of the variables are both dynamic and interactive. The task-based account may thus promote greater compatibility with neurological evidence suggestive of a tight linkage between category symbolization and action representation, cf. the emerging evidence on bimodal and mirror neurons (e.g., Gallese and Goldman, 1998, Rizzolati *et al.* 2001; Maravita and Iriki, 2004).

What is representation for?

The classical, symbolic view promotes the idea that the purpose of representation is to provide an internal locus of contingencies relating to external entities and events. Non-classical approaches, in contrast, tend to view representational structure as embedded within other types of processing and have the potential to identify the purpose of representation with non-representational function. Clark and Grush, for example, see representation emerging in 'emulator circuits', where the immediate goal is the provision (through simulation) of the rapid sensory feedback which is vital for smooth motor-control. The purpose of representation in that context can be viewed as the improvement of 'real-world, real-time responsiveness' (Clark and Grush, 1999, p. 7). Contrast this with the task-analytic answer to the question, which is that the purpose of representation can be seen as gain of information. Though different, the two answers are not incompatible, however.

Introduction of encapsulative structure for purposes of information-gain may be seen as a way of dealing with the information *losses* which result from distributive effects. Encapsulative representation may then be viewed as a means of *recovering* information losses embodied in the distal-to-proximal mapping. The effective aim is then to establish a tight, informational linkage with distal stimulation, i.e., to establish what Clark and Grush term 'real-world, real-time responsiveness'. Their view of what representation is for thus chimes reasonably well with that emerging from the task analysis.

What the analysis can add, however, is a formal notion of how the strategy of linkage-making (through representation) can be traded-off with the strategy which eliminates the information-losing 'separation'. Figure 16 provides a simple illustration. At the top of the figure, we have a distal sensing scheme embodying a single intermediate variable. The distributive (i.e., 'mixed-up') nature of the target's influence on this variable produces an information loss which intra-variable reversal cannot fully remedy. Contrasting the sensor gain of 0.161 bbits with the maximum of $log_2(3) = 1.58$ bbits, the size of this loss can be measured at 1.42 bbits.

In the lower part of the figure we then see schemes representing alternative ways in which the loss can be dealt with. In the scheme on the left, the loss is *eliminated* by the expedient of removing the distributive structure altogether. Target T becomes a receptor, enabling conditional relationships to connect values of T directly with the appropriate values of S. Sensor and target are brought into a closely connected relationship, effectively providing proximal access to the distal stimulus.

In the scheme on the right, an encapsulative variable is introduced along with an additional intermediate variable. Exploiting constraints operating across the

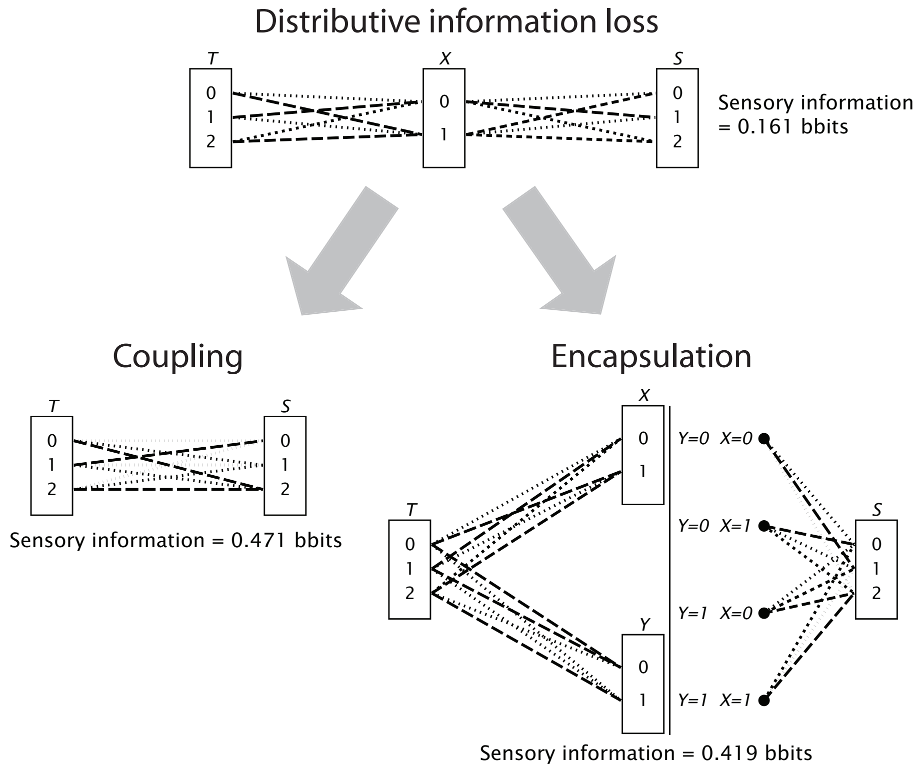

Figure 16: Recovery of lost information through encapsulation.

two intermediate variables, this then provides the means of *recovering* the lost information, with the achieved level of performance coming close to that of the closely coupled system.

The example illustrates how introduction of encapsulation can be the means of recovering lost information. It also illustrates the complementarity between representation-use and *coupling*. Again, this meshes with Clark and Grush (1999), who note that representation-use can be understood as a vehicle of *de*-coupling. The specific implication arising from the task analysis, however, is that the representation/coupling relationship takes the form of a tradeoff. Where distal-to-proximal seperation engenders loss of information, the deficit may be resolved either by information recovery (by use of representation) or through elimination of the seperation. Relatively greater use of one strategy allows for relatively less of the other.

How can representation be acquired?

On the classical, symbolic view, the acquisition of representational structure seems to necessitate a logically prior step involving the acquisition of the language which gives meaning to symbols (Gärdenfors, 2000, pp. 37-40). Non-classical approaches dispense with the prior requirement for an encoding language, replacing it with a requirement for whatever is prerequisite for execution of the function in which representation is embedded. This might mean acquisition of real-time sensory functionality of some sort. But the task-analytic suggestion is that acquisition of representation ultimately entails discovery of those invariant properties which underlie salient distributive effects.

This process might proceed on the basis of the detection of mutual-information between sensory streams, as envisaged by (Pfeifer and Scheier, 1999), or on the basis of information self-structuring (Lungarella and Sporns, 2005), or on the basis of some kind of information pickup, such as the 'atunement' operation envisaged by Gibson (1979). Indeed, the process Gibson envisages seems particularly relevant to discovery of invariants underlying distributive effects. He defines 'atunement' as the 'registering of certain definite dimensions of invariance in the stimulus flux along with definite parameters of disturbance.' (Gibson, 1979, p. 249) He also conceptualizes the nature of those invariances in a way which is compatible with their being informational. As he notes 'Mathematical complexities of stimulus energy seem to be the simplicities of stimulus information.' (Gibson, 1966, p. 319)

When is representation unnecessary?

Just as the symbolic view encounters circularities in explaining why representation is necessary, so it suffers problems in accounting for those scenarios where representation is clearly *unnecessary* (e.g., Brooks, 1986; Maes and Brooks, 1990; Beer, 1990; Pfeifer and Scheier, 1997). A well-known illustration of such a scenario is the cross-connected version of Braitenberg's 'vehicle 2' robot (Braitenberg, 1984): see schematic (b) in Figure 16. This robot executes a robust pursuit behavior with regard to a sensory target.

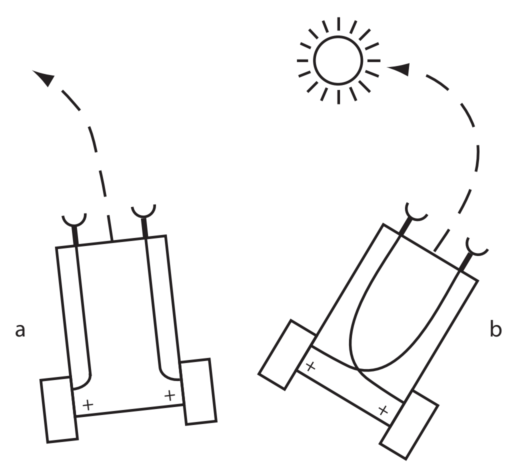

Figure 17: Braitenberg's vehicle 2 architectures (Braitenberg, 1984, p.8).

The target is taken to directly stimulate the robot's two receptors with an intensity that is proportional to distance, and with a differential value that depends on orientation. With stimulation to each receptor transmitted to the wheel motor on the opposite side of the robot, a sensory-motor linkage is obtained whereby the robot always turns/moves towards the source of stimulation. The symbolic view of representation seems to suggest no particular explanation as to why this robust behavior is achievable without use of representational structure. But the task-analytic view is more specific: representation is not required in this scenario because there is *no loss of sensory information*.

Sensing in this context can be understood to involve the impact on two receptors produced by two variables, namely 'leftside-stimulation' and 'rightside-stimulation'. With this 1-to-1 mapping, there is no distributive structure in the environment and no information loss. Recovery of information is not required.

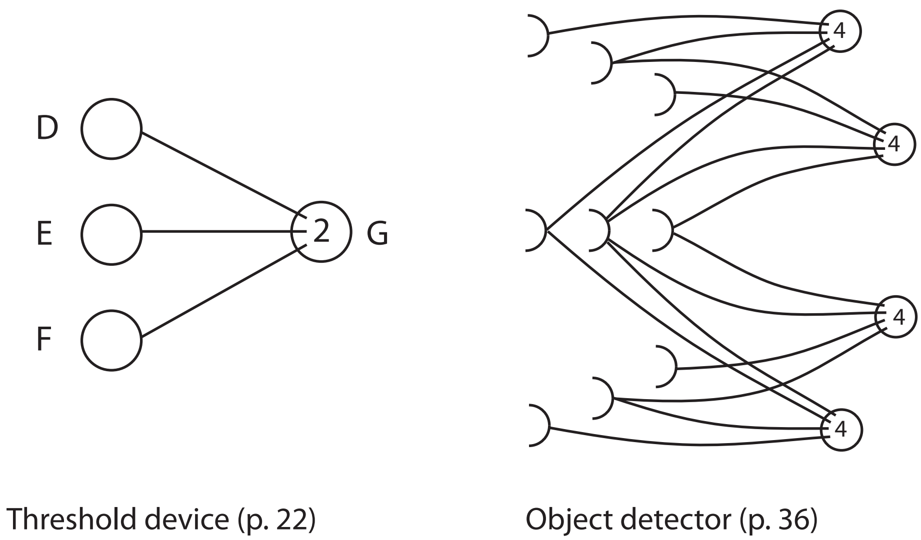

Figure 18: Encapsulative devices from (Braitenberg, 1984).

In his vehicle 2 architecture, Braitenberg pursues a minimalist style of implementation. What makes the architecture interesting for present purposes, however, is the fact that it represents an extreme point on the representation/coupling tradeoff. Behavioral efficacy is achieved solely on the basis of coupling. Representation plays no role. However, it is easy to envisage how a small adjustment in the representation/coupling balance might be made.

On the present analysis, any representation-increasing (coupling-reducing) adjustment must involve introduction of some element of encapsulative structure. We can think of this in terms of a simple distributive/encapsulative scheme (along the lines of Figure 7 or Figure 12) being interposed between sensors and motors. But we can also make use of Braitenberg's own proposals. In his 'brand 5 vehicles' (p. 20-25) use is made of what Braitenberg calls 'threshold devices', see Figure 18.[16] While these operate in the manner of simple artificial neurons, the fact that they set their state according to a combinational property of their inputs means they can be interpreted as encapsulative variables. Braitenberg shows how their use can be the basis of sensitivity to a range of distal stimuli including ‘an olive green vehicle that buzzes at a certain frequency and never moves faster than 5m/sec’ (p. 22). This is the effect we expect the introduction of encapsulative representational structure to have.

In his Vehicle 8 designs, Braitenberg explores how complex networks of encapsulative structures — e.g., the object detector of Figure 18 — can serve to provide the basis for detection of spatially-organized distal stimuli, i.e., locations in a map. These examples illustrate how the simple distributive/encapsulative schemes considered in the present paper might scale-up. They highlight some of the ways in which the assembly of encapsulative structure into complexes can be the basis for scene recognition and visual interpretation. In sum, they illustrate how a primitive information-recovery mechanism, when combined, cross-connected and cascaded in the right ways, can serve to provide the forms of recognizably *symbolic* representation with which we are most familiar.

8 Concluding comment

It is now more than 20 years since Brooks proposed that representation-use can create an ‘information bottleneck’ in sensory-motor systems (Brooks, 1986). His thesis surely remains correct with regard to the usage he envisaged, i.e., modular programs mediating multi-stage cross-referencing between world models and sequential processing streams. That approach, however, was essentially a wholesale import from computer science, a field with very different concerns.[17] Notions of representation-use have now moved on in significant ways.

Guided by novel, non-symbolic conceptions, researchers are beginning to uncover forms of sensorily-embedded representation which rely on interconnected surrogacies and system-functional content. The present paper shows, on purely theoretical grounds, why we should expect utilization of such forms to be widespread in nature. They are the means of achieving efficiency in distal sensing and are therefore likely to be critical for interpretive function. Brooks was right to say that representation-usage has significant informational consequences but he may have been wrong in describing what those consequences are. Rather than serving to obstruct information, representations are better seen as the means of regenerating or recovering it.

9 Appendix 1

[16] The threshold device is the central elements in Braitenberg’s Figure 9. The object detector forms Figure 12.

[17] Stand-alone representational structures have well understood advantages in computer science (Sedgewick, 1988).

Conditional probabilities for green lacewing moth models		
Receptor values	Informative sensor	Misinformative sensor
P(FM=0|T=0) = 0.1 P(FM=1|T=0) = 0.9 P(FM=0|T=1) = 0.9 P(FM=1|T=1) = 0.1 P(PF=0|T=0) = 0.9 P(PF=1|T=0) = 0.1 P(PF=0|T=1) = 0.1 P(PF=1|T=1) = 0.9	P(S=0|FM=0,PF=0) = 0.5 P(S=1|FM=0,PF=0) = 0.5 P(S=0|FM=0,PF=1) = 0.02 P(S=1|FM=0,PF=1) = 0.98 P(S=0|FM=1,PF=0) = 0.98 P(S=1|FM=1,PF=0) = 0.02 P(S=0|FM=1,PF=1) = 0.5 P(S=1|FM=1,PF=1) = 0.5	P(S=0|FM=0,PF=0) = 0.5 P(S=1|FM=0,PF=0) = 0.5 P(S=0|FM=0,PF=1) = 0.84 P(S=1|FM=0,PF=1) = 0.16 P(S=0|FM=1,PF=0) = 0.16 P(S=1|FM=1,PF=0) = 0.84 P(S=0|FM=1,PF=1) = 0.5 P(S=1|FM=1,PF=1) = 0.5

References

Ballard, D. (1991). Animate vision. *Artificial Intelligence*, *48* (pp. 57-86).

Barlow, H. (1961). Possible principles underlying the transformation of sensory messages. In W.A. Rosenblith (Ed.), *Sensory communication* (pp. 217-234). Cambridge: MIT Press.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22* (pp. 577-660).

Bateson, G. (2000). *Steps to an Ecology of Mind*. Chicago: University of Chicago Press.

Beer, R. (1990). *Intelligence as Adaptive Behaviour: An Experiment in Computational Neuroethology*. San Diego: Academic Press.

Beer, R. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*, No. 4 (pp. 209-243).

Bickhard, M. and Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. New York: Elsevier Science.

Brachman, R. and Levesque, H. (Eds.) (1985). *Readings in Knowledge Representation*. Los Altos: Morgan Kaufmann.

Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. London: The MIT Press.

Brignell, J. and White, N. (1994). *Intelligent Sensor Systems*. Bristol: IOP Publishing Ltd.

Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, *RA-2*, No. 1 (pp. 14-23).

Brooks, R. (1991a). Intelligence without reason. In J. Mylopoulos and R. Reiter (Eds.), *Proceedings of the Twelth International Joint Conference on Artificial Intelligence* (pp. 569-595). San Mateo, California: Morgan Kaufman.

Brooks, R. (1991b). Intelligence without representation. *Artificial Intelligence*, *47* (pp. 139-159).

Clancey, W. (1997). *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge: Cambridge University Press.

Clark, A. (1996). Happy couplings: emergence and explanatory interlock. In M.A. Boden (Ed.), *The Philosophy of Artificial Life* (pp. 262-281). Oxford: Oxford University Press.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

Clark, A. (2003). *Natural Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*. Oxford University Press.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action and Cognitive Extension*. Oxford: Oxford University Press.

Clark, A. and Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, *7*, No. 1 (pp. 5-16).

Clark, A. and Thornton, C. (1997). Trading spaces: computation, representation and the limits of uninformed learning. *Behaviour and Brain Sciences*, *20* (pp. 57-90). Cambridge University Press.

Clark, A. and Toribio, J. (1994). Doing without representing. *Synthese*, *101* (pp. 401-431).

Cover, T. and Thomas, J. (1991). *Elements of Information Theory*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Dean, P., Mayhew, J. and Langdon, P. (1994). Learning and maintaining saccadic accuracy: a model of brainstem-cerebellum interactions. *Jounral of Cognitive Neuroscience*, *6* (pp. 117-138).

Dreyfus, H. (1979). *What Computers Cant́ Do* (revised edition). New York: Harper and Row.

Fodor, J. (1975). *The Language of Thought*. New York: Crowell.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, No. 12 (pp. 493-501).

Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. Cambridge, Mass.: The MIT Press.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Oxford, England: Houghton-Mifflin.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Miffin.

Grush, R. (2004). The emulation theory of representation. *Behavioral and Brain Sciences*, *27* (pp. 377-442).

Harnad, S. (1990). The symbol grounding problem. *Physica D*, *42* (pp. 335-346).

Haselager, P., de Groot, A. and van Rappard, H. (2003). Representationalism vs. anti-representationalism: a debate for the sake of appearance. *Philosophical Psychology*, *16* (pp. 5-23).

Haugeland, J. (1991). Representational genera. In W. Ramsay, S.P. Stich and D.E. Rumelhart (Eds.), *Philosophy and Connectionist Theory* (pp. 61-89). Mahwah, NJ: Lawrence Erlbaum Associates.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, *6* (pp. 242-247).

Hillis, J., Ernst, M., Banks, M. and Landy, M. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science*, *298* (pp. 1627-1630).

Kawato, M., Furukawa, K. and Suzuki, R. (1987). A hierarchical neural network model for the control and learning of voluntary movement. *Biological Cybernetics*, *57* (pp. 169-185).

Kirsh, D. (1996). Today the earwig, tomorrow man?. In M.A. Boden (Ed.), *The Philosophy of Artificial Life* (pp. 237-261). Oxford: Oxford University Press.

Korb, K. and Nicholson, A. (2003). *Bayesian Artificial Intelligence*. Boca Raton, Florida: CRC Press.

Kullback, S. (1987). The kullback-leibler distance. *The American Statistician*, *41* (pp. 340-341).

Laming, D. (1986). *Sensory Analysis*. London: Academic Press.

Lungarella, M., Pegors, T., Bulwinkle, D. and Sporns, O. (2005). Methods for quantifying the information structure of sensory and motor data. *Neuroinformatics*, *3*, No. 3 (pp. 243-2620).

Lungarella, M. and Pfeifer, R. (2001). Robots as cognitive tools: information theoretic analysis of sensory-motor data. *Proceedings of the 2001 IEEE-RAS International Conference on Humanoid Robots* (pp. 245-252). Piscataway, NJ: IEEE Press.

Lungarella, M. and Sporns, O. (2005). Information self-structuring: key principles for learning and development. *Proceedings 2005 4th IEEE International Conference on Development and Learning and Learning* (pp. 25-30).

Maes, P. and Brooks, R. (1990). Learning to coordinate behaviors. *Proceedings of the Eighth National Conference on Artificial Intelligence* (pp. 796-802). San Francisco, CA: Morgan Kaufmann.

Maravita, A. and Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, *8*, No. 2 (pp. 79-86).

Marr, D. (1982). *Vision*. New York: W.H. Freeman.

Mather, P. (1999). *Computer Processing of Remotely-Sensed Images*. Chichester, England: John Wiley and Sons Ltd.

Newell, A. and Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.

Newell, A. and Simon, H. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, *19*, No. 3 (pp. 113-126).

Noë, A. (2004). *Action in Perception*. London, England: The MIT Press.

Pearl, J. (1985). Bayesian networks: a model of self-activated memory for evidential reasoning. *Proceedings of the 7th Conference of the Cognitive Science Society* (pp. 329-334). Irvine, CA: University of California, Irvine, CA.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge, England: Cambridge University Press.

Pfeifer, R. and Scheier, C. (1997). Sensory-motor coordination: the metaphor and beyond. practice and future of autonomous agents. *Robotics and Autonomous Systems*, *20* (pp. 157-178).

Pfeifer, R. and Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Prinz, J. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA: MIT Press.

Pylyshyn, Z. (1987). *The Robots Dilemma: The Frame Problem in Artificial Intelligence*. Norwood: Ablex.

Rizzolati, G., Forgassi, L. and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews: Neuroscience*, *2* (pp. 661-670).

Sabins, F. (1978). *Remote Sensing: Principles and Interpretation*. New York: W.H. Freeman and Company.

Searle, J. (1980). Minds, brains and programs [with peer commentaries]. *Behavioural and Brain Sciences*, *3* (pp. 417-57).

Sedgewick, R. (1988). *Algorithmics* (2nd edition). Menlo Park, California: Addison-Wesley Publising Company, Inc.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27* (pp. 379-423 and 623-656).

Shannon, C. and Weaver, W. (1949). *The Mathematical Theory of Communication.* Urbana, Illinois: University of Illinois Press.

Smith, B. (1996). *On the Origin of Objects.* Cambridge, MA: MIT Press.

Smith, C. (2000). *Biology of Sensory Systems.* New York: John Wiley & Sons, Ltd.

Stewart, J. (1995). Cognition = life: implications for higher-level cognition. *Behavioural Processes*, *35* (pp. 311-326).

Svensson, H. and Ziemke, T. (2005). Embodied representation: what are the issues?. In B.G. Bara, L. Barsalou and M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 2116-2121). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tononi, G., Sporns, O. and Edelman, G. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the Nat. Academy of Science*, *91* (pp. 5033-5037).

Tononi, G., Sporns, O. and Edelman, G. (1996). A complexity measure for selective matching of signals by the brain. *Proceedings of the Nat. Academy of Science*, *93* (pp. 3422-3427).

Usher, M. and Keating, D. (1996). *Sensors and Transducers: Characteristics, Applications, Instrumentation, Interfacing.* London: Macmillan Press Ltd.

van Gelder, T. (1995). What might cognition be, if not computation. *Journal of Philosophy*, *XCII*, No. 7 (pp. 345-381).

van Hateren, J. (1992). A theory of maximizing sensory information. *Biological Cybernetics*, *68*, No. 1 (pp. 23-29). Berlin/Heidelberg: Springer.

Varella, F., Thompson, E. and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience.* Cambridge, Massachusetts: The MIT Press.

Webster, J. (1999). *The Measurement, Instrumentation and Sensors Handbook.* Danvers, MA: CRC Press LLC.

Wheeler, M. (1996). From robots to rothko: the bringing forth of worlds. In M.A. Boden (Ed.), *The Philosophy of Artificial Life* (pp. 209-236). Oxford: Oxford University Press.

Wheeler, M. and Clark, A. (1999). Genic representation: reconciling content and causal complexity. *The British Journal for the Philosophy of Science*, *50* (pp. 103-135). Br Soc Philosophy Sci.

Wilson, J. (2005). *Sensor Technology Handbook*. Oxford: Elsevier Inc.

Winograd, T. and Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, New Jersea: Ablex Publishing Corporation.